time stamp (PTS) and the decode time stamp (DTS). The audio and video sampling clocks, f_a and f_v respectively, must be frequency-locked to the 27 MHz clock. This can be expressed as the requirement that there exist two pairs of integers, (n_a, m_a) and (n_v, m_v), such that:

$$f_a = \left(\frac{n_a}{m_a}\right) x\ 27MHz$$

and

$$f_v = \left(\frac{n_v}{m_v}\right) x\ 27MHz$$

The channel coding domain is represented by the FEC/Sync Insertion subsystem and the VSB modulator. The relevant frequencies in this domain are the VSB symbol frequency (f_{sym})and the frequency of the transport stream (f_{TP}) which is the frequency of transmission of the encoded transport stream. These two frequencies must be locked, having the relation:

$$f_{TP} = 2\ x \left(\frac{188}{208}\right)\left(\frac{312}{313}\right) f_{sym}$$

The signals in the two domains are not required to be frequency-locked to each other, and in many implementations will operate asynchronously. In such systems, the frequency drift can necessitate the occasional insertion or deletion of a NULL packet from within the transport stream, thereby accommodating the frequency disparity.

The annexes that follow consider the characteristics of the subsystems necessary to accommodate the services envisioned.

ANNEX A

(Normative)

VIDEO SYSTEM CHARACTERISTICS

Table of Contents

List of Tables

Blank Page

ANNEX A

(Normative)

VIDEO SYSTEM CHARACTERISTICS

1. SCOPE

This Annex describes the characteristics of the video subsystem of the Digital Television Standard. The input formats and bit stream characteristics are described in separate sections.

2. REFERENCES

2.1 Normative references

The following documents contain provisions which, through reference in this text, constitute provisions of this standard. At the time of publication, the editions indicated were valid. All standards are subject to revision, and parties to agreement based on this standard are encouraged to investigate the possibility of applying the most recent editions of the documents listed below.

ISO/IEC IS 13818-1, International Standard (1994), *MPEG-2 Systems.*

ISO/IEC IS 13818-2, International Standard (1994), *MPEG-2 Video.*

2.2 Informative references

SMPTE S17.392 (1995), *Proposed Standard for television, 1280 x 720 Scanning and Interface.*

SMPTE 274M (1994), *Standard for television, 1920 x 1080 Scanning and Interface.*

3. COMPLIANCE NOTATION

As used in this document, *"shall"* or *"will"* denotes a mandatory provision of the standard. *"Should"* denotes a provision that is recommended but not mandatory. *"May"* denotes a feature whose presence does not preclude compliance, that may or may not be present at the option of the implementor.

4. VIDEO INPUT FORMATS

Two HDTV production formats, SMPTE S17.392 and SMPTE 274M, may be accepted as inputs to the video subsystem.[5] The main parameters of these formats, given in Table 1, closely relate to those of the compression formats (see Table 3).

Table 1 Video Input Formats

Active samples/ line	Active lines	Image rate[1] (Hz)	Progressive/ interlaced	Aspect ratio information
1280	720	60.00, 59.94	progressive	16:9 square pixels
1920	1080	60.00, 59.94	progressive	16:9 square pixels
1920	1080	60.00, 59.94	interlaced	16:9 square pixels

5. SOURCE CODING SPECIFICATION

The ATV video compression algorithm shall conform to the Main Profile syntax of ISO/IEC 13818-2. The allowable parameters shall be bounded by the upper limits specified for the Main Profile at High Level.[6] Additionally, ATV bit streams shall meet the constraints and specifications described in Sections 5.1 and 5.2.

An ATV decoder will decode the subset of MP@HL bit streams which meet the constraints found in Section 5.1. A decoder which conforms to ISO 13818-2 MP@HL is expected to properly decode all ATV bit streams.[7]

5.1 Constraints with respect to MPEG-2 MP@HL

The following tables list the allowed values for each of the ISO 13818-2 syntactic elements which are restricted beyond the limits imposed by MP@HL.

In these tables conventional numbers denote decimal values, numbers preceded by **0x** are to be interpreted as hexadecimal values and numbers within single quotes (e.g., '10010100') are to be interpreted as a string of binary digits.

5.1.1 Sequence header constraints

Table 2 identifies parameters in the sequence header of a bit stream that shall be constrained by the video subsystem and lists the allowed values for each.

---

[5] The SMPTE formats in their present form provide for only 60.00/59.94 Hz image rates. The video subsystem may also accept as inputs HDTV production formats at 24.00/23.98 and 30.00/29.97 Hz that otherwise conform to the spatial and other parameters of the SMPTE formats.

[6] See ISO/IEC 13818-2, Section 8 for more information regarding profiles and levels.

[7] In order to decode the user data, the decoder should properly recognize the 32-bit ATSC registration identifier at the PSI stream level (see ISO/IEC 13818-1).

Table 2 Sequence Header Constraints

Sequence header syntactic element	Allowed value
horizontal_size_value	see Table 3
vertical_size_value	see Table 3
aspect_ratio_information	see Table 3
frame_rate_code	see Table 3
bit_rate_value (8 VSB)	≤ 48500
bit_rate_value (16 VSB)	≤ 97000
vbv_buffer_size_value	≤ 488

The allowable values for the field bit_rate_value are application dependent. In the primary application of terrestrial broadcast, this field shall correspond to a bit rate which is less than or equal to 19.4 Mbps. In the case of cable systems using the high data rate cable mode of 16 VSB transmission, the corresponding bit rate shall be less than or equal to 38.8 Mbps.

5.1.2 Compression format constraints

Table 3 lists the allowed compression formats. They correspond to the supported video input formats shown in Table 1.

Table 3 Compression Format Constraints

horizontal_size_ value	vertical_size_ value	frame_rate_ code	progressive_ sequence	aspect_ratio_ information
1280	720	1,2,4,5,7,8	1	1,3
1920	1080[8]	1,2,4,5	1	1,3
1920	1080[4]	4,5	0	1,3

5.1.3 Sequence extension constraints

Table 4 identifies parameters in the sequence extension part of a bit stream that shall be constrained by the video subsystem and lists the allowed values for each. A sequence_extension structure is required to be present after every sequence_header structure.

Table 4 Sequence Extension Constraints

Sequence extension syntactic element	Allowed values
progressive_sequence	see Table 3
profile_and_level_indication	'11000100'
chroma_format	'01'
horizontal_size_extension	'00'
vertical_size_extension	'00'

[8] Note that 1088 lines are actually coded in order to satisfy the MPEG-2 requirement that the coded vertical size be a multiple of 32.

Sequence extension syntactic element	Allowed values
bit_rate_extension	'000000000000'
vbv_buffer_size_extension	'00000000'
frame_rate_extension_n	'00'
frame_rate_extension_d	'00000'

5.1.4 Sequence display extension constraints

Table 5 identifies parameters in the sequence display extension part of a bit stream that shall be constrained by the video subsystem and lists the allowed values for each.

Table 5 Sequence Display Extension Constraints

Sequence display extension syntactic element	Allowed values
video_format	'000'
color_primaries	0x07
transfer_characteristics	0x07
matrix_coefficients	0x07

5.1.5 Picture header constraints

In all cases other than when vbv_delay has the value 0xFFFF, the value of vbv_delay shall be constrained as follows:

$$vbv\_delay \leq 45000$$

5.2 Bit stream specifications beyond MPEG-2

This section covers the extension and user data part of the video syntax. These data are inserted at the sequence, GOP, and picture level. The syntax used for the insertion of closed captioning in picture user data is described.

5.2.1 Picture extension and user data syntax

Table 6 describes the syntax used for picture extension and user data.

Table 6 Picture Extension and User Data Syntax

	No. of bits	Mnemonic
extension_and_user_data(2) {		
while ((nextbits( ) == extension_start_code) || (nextbits() == user_data_start_code)) {		
if (nextbits()== extension_start_code)		
extension_data(2)		
if (nextbits() == user_data_start_code)		
user_data(2)		
}		
}		

5.2.2 Picture user data syntax

Table 7 describes the picture user data syntax.

Table 7 Picture User Data Syntax[9]

	No. of bits	Mnemonic
user_data( ) {		
user_data_start_code	32	bslbf
ATSC_identifier	32	bslbf
user_data_type_code	8	uimsbf
if (user_data_type_code == '0x03') {		
process_em_data_flag	1	bslbf
process_cc_data_flag	1	bslbf
additional_data_flag	1	bslbf
em_data	8	bslbf
cc_count	5	uimsbf
for (i=0 ; i < cc_count ; i++) {		
marker_bit	1	bslbf
cc_priority	2	uimsbf
cc_data_1	8	bslbf
cc_data_2	8	bslbf
}		
marker_bit	1	bslbf
while (!bytealigned( )) {		
marker_bit	1	bslbf
}		
if (additional_data_flag) {		
while(nextbits() != '0000 0000 0000 0000 0000 0001') {		
additional_user_data	8	
}		
}		
}		
next_start_code()		
}		

5.2.3 Picture user data semantics

user_data_start_code — This is set to 0x0000 01B2.

ATSC_identifier — This is a 32 bit code that indicates that the video user data conforms to this specification. The value ATSC_identifier shall be 0x4741 3934.

user_data_type_code — The 8-bit code is set to 0x03.

[9] Shaded cells in this table indicate syntactic and semantic additions to the ISO/IEC 13818-2 standard.

process_em_data_flag — This flag is set to indicate whether it is necessary to process the em_data. If it is set to 1, the em_data has to be parsed and its meaning has to be processed. When it is set to 0, the em_data can be discarded.

process_cc_data_flag — This flag is set to indicate whether it is necessary to process the cc_data. If it is set to 1, the cc_data has to be parsed and its meaning has to be processed. When it is set to 0, the cc_data can be discarded.

additional_data_flag — This flag is set to 1 to indicate the presence of additional user data.

em_data — Eight bits for representing emergency message.[10]

cc_count — This 5-bit integer indicates the number of closed caption constructs following this field. It can have values 0 through 31. All such constructs must occur in the intended display order, assuming interlaced display. The sum of cc_count for all pictures in any one-second interval shall be equal to 600.

cc_priority — A number between 0 and 3 used to indicate the priority of constructs in picture reconstruction where different levels of hardware capability exist. For closed caption constructs, up to four lines per display field can be labeled as priority zero.

cc_data_1 — The first byte of a closed caption data pair.[11]

cc_data_2 — The second byte of a closed caption data pair.[7]

bytealigned() — This function is defined in ISO/IEC 13818-2. This function returns 1 if the current position is on a byte boundary, that is, the next bit in the bit stream is the first bit in a byte. Otherwise it returns 0.

additional_user_data — Any further demand for picture user data could be met by defining this part of the bit stream.

---

[10] Syntax and semantics to be specified by EIA.

[11] EIA, *Recommended Practice for Advanced Television Closed Captioning*, draft, July 1, 1994.

ANNEX B

(Normative)

AUDIO SYSTEM CHARACTERISTICS

Table of Contents

List of Figures

List of Tables

ANNEX B

(Normative)

AUDIO SYSTEM CHARACTERISTICS

1. SCOPE

This Annex describes the audio system characteristics and normative specifications of the Digital Television Standard.

2. NORMATIVE REFERENCES

The following documents contain provisions which in whole or part, through reference in this text, constitute provisions of this standard. At the time of publication, the editions indicated were valid. All standards are subject to revision and amendment, and parties to agreement based on this standard are encouraged to investigate the possibility of applying the most recent editions of the documents listed below.

ATSC Standard A/52 (1994), *Digital Audio Compression (AC-3).*

AES 3-1992 (ANSI S4.40-1992), *AES Recommended Practice for digital audio engineering — Serial transmission format for two-channel linearly represented digital audio data.*

ANSI S1.4-1983, *Specification for Sound Level Meters.*

IEC 651 (1979), *Sound Level Meters.*

IEC 804 (1985), Amendment 1 (1989) *Integrating/Averaging Sound Level Meters.*

3. COMPLIANCE NOTATION

As used in this document, *"shall"* or *"will"* denotes a mandatory provision of the standard. *"Should"* denotes a provision that is recommended but not mandatory. *"May"* denotes a feature whose presence does not preclude compliance, that may or may not be present at the option of the implementor.

4. SYSTEM OVERVIEW

As illustrated in Figure 1, the audio subsystem comprises the audio encoding/decoding function and resides between the audio inputs/outputs and the transport subsystem. The audio encoder(s) is (are) responsible for generating the audio elementary stream(s) which are encoded representations of the baseband audio input signals. At the receiver, the audio subsystem is responsible for decoding the audio elementary stream(s) back into baseband audio.

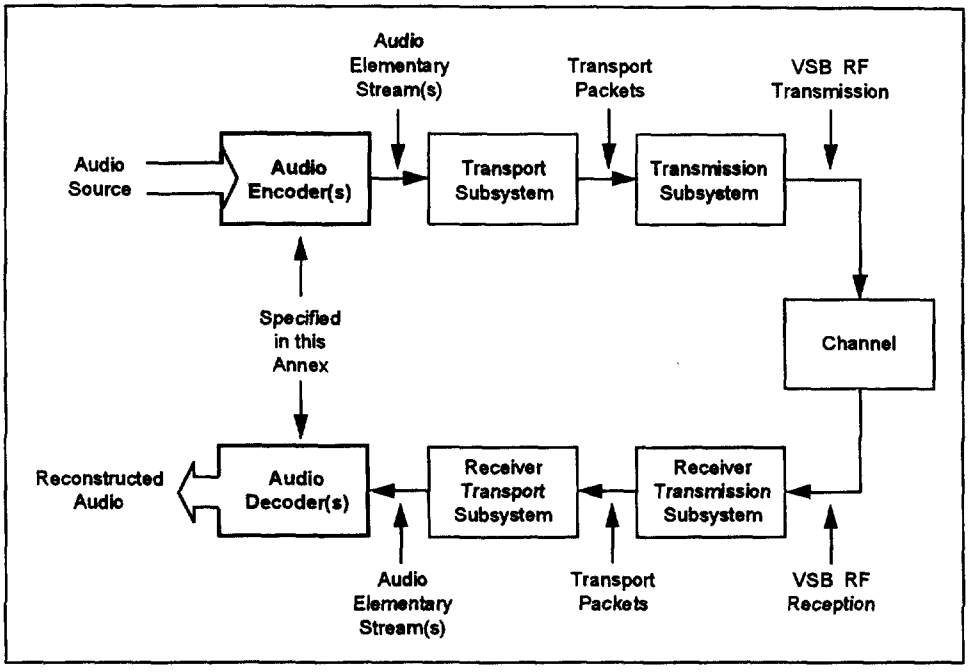

Figure 1. Audio subsystem in the digital television system.

5. SPECIFICATION

This Section forms the normative specification of the audio system. The audio compression system conforms with the Digital Audio Compression (AC-3) Standard, subject to the constraints outlined in this Section.

5.1 Constraints with respect to ATSC Standard A/52

The digital television audio coding system is based on the Digital Audio Compression (AC-3) Standard specified in ATSC Doc. A/52. Constraints on the system are shown in Table 1 which shows permitted values of certain syntactical elements. These constraints are described in Sections 5.2 - 5.4.

Table 1 Audio Constraints

AC-3 syntactical element	Comment	Allowed value
fscod	Indicates sampling rate	00 (indicates 48 kHz)
frmsizecod	Main audio service	≤ 011100 (indicates ≤ 384 kbps)
frmsizecod	Single channel associated service	≤ 010000 (indicates ≤ 128 kbps)
frmsizecod	Two channel associated service	≤ 010100 (indicates ≤ 192 kbps)
(frmsizecod)	Combined bit rate of a main and an associated service intended to be simultaneously decoded	(total ≤ 512 kbps)
acmod	Indicates number of channels	≥ 001

5.2 Sampling frequency

The system conveys digital audio sampled at a frequency of 48 kHz, locked to the 27 MHz system clock. The 48 kHz audio sampling clock is defined as:

(1) 48 kHz audio sample rate = (2 ÷1125) × (27 MHz system clock)

If analog signal inputs are employed, the A/D converters should sample at 48 kHz. If digital inputs are employed, the input sampling rate shall be 48 kHz, or the audio encoder shall contain sampling rate converters which convert the sampling rate to 48 kHz.

5.3 Bit rate

A main audio service shall be encoded at a bit rate less than or equal to 384 kbps. A single channel associated service shall be encoded at a bit rate less than or equal to 128 kbps. A two channel associated service shall be encoded at a bit rate less than or equal to 192 kbps. The combined bit rate of a main service and an associated service which are intended to be decoded simultaneously shall be less than or equal to 512 kbps.

5.4 Audio coding modes

Audio services shall be encoded using any of the audio coding modes specified in A/52, with the exception of the 1+1 mode. The value of acmod in the AC-3 bit stream shall have a value in the range of 1-7, with the value 0 prohibited.

5.5 Dialogue level

The value of the dialnorm parameter in the AC-3 elementary bit stream shall indicate the level of average spoken dialogue within the encoded audio program. Dialogue level may be measured by means of an "A" weighted integrated measurement (L_{Aeq}). (Receivers use the value of dialnorm to adjust the reproduced audio level so as to normalize the dialogue level.)

5.6 Dynamic range compression

Each encoded audio block may contain a dynamic range control word (dynrng) which is used by decoders (by default) to alter the level of the reproduced audio. The control words allow the decoded signal level to be increased or decreased by up to 24 dB. In general, elementary streams may have dynamic range control words inserted or modified without affecting the encoded audio. When it is necessary to alter the dynamic range of audio programs which are broadcast, the dynamic range control word should be used.

6. MAIN AND ASSOCIATED SERVICES

6.1 Overview

An AC-3 elementary stream contains the encoded representation of a single audio service. Multiple audio services are provided by multiple elementary streams. Each

elementary stream is conveyed by the transport multiplex with a unique PID. There are a number of audio service types which may (individually) be coded into each elementary stream. Each AC-3 elementary stream is tagged as to its service type using the bsmod bit field. There are two types of *main service* and six types of *associated service*. Each associated service may be tagged (in the AC-3 audio descriptor in the transport PSI data) as being associated with one or more main audio services. Each AC-3 elementary stream may also be tagged with a language code.

This Section specifies the meaning and use of each type of service. In general, a complete audio program (what is presented to the listener over the set of loudspeakers) may consist of a main audio service, or a main audio service combined with an associated audio service. The capability to simultaneously decode one main service and one associated service is required in order to form a complete audio program in certain service combinations described in this Section.

6.2 Summary of service types

The audio service types are listed in Table 2.

Table 2 Audio Service Types

bsmod	Type of service
000 (0)	Main audio service: complete main (CM)
001 (1)	Main audio service: music and effects (ME)
010 (2)	Associated service: visually impaired (VI)
011 (3)	Associated service: hearing impaired (HI)
100 (4)	Associated service: dialogue (D)
101 (5)	Associated service: commentary (C)
110 (6)	Associated service: emergency (E)
111 (7)	Associated service: voice-over (VO)

6.3 Complete main audio service (CM)

The CM type of main audio service contains a complete audio program (complete with dialogue, music, and effects). This is the type of audio service normally provided. The CM service may contain from 1 to 5.1 audio channels. The CM service may be further enhanced by means of the VI, HI, C, E, or VO associated services described below. Audio in multiple languages may be provided by supplying multiple CM services, each in a different language.

6.4 Main audio service, music and effects (ME)

The ME type of main audio service contains the music and effects of an audio program, but not the dialogue for the program. The ME service may contain from 1 to 5.1 audio channels. The primary program dialogue is missing and (if any exists) is supplied by simultaneously encoding a D service. Multiple D services in different languages may be associated with a single ME service.

6.5 Visually impaired (VI)

The VI associated service typically contains a narrative description of the visual program content. In this case, the VI service shall be a single audio channel. The simultaneous reproduction of both the VI associated service and the CM main audio service allows the visually impaired user to enjoy the main multi-channel audio program, as well as to follow (by ear) the on-screen activity.

The dynamic range control signal in the VI service is intended to be used by the audio decoder to modify the level of the main audio program. Thus the level of the main audio service will be under the control of the VI service provider, and the provider may signal the decoder (by altering the dynamic range control words embedded in the VI audio elementary stream) to reduce the level of the main audio service by up to 24 dB in order to assure that the narrative description is intelligible.

Besides providing the VI service as a single narrative channel, the VI service may be provided as a complete program mix containing music, effects, dialogue, and the narration. In this case, the service may be coded using any number of channels (up to 5.1). The fact that the service is a complete mix shall be indicated in the AC-3 descriptor (see Section 5.7.2.1 in Annex C of this Standard).

6.6 Hearing impaired (HI)

The HI associated service typically contains only dialogue which is intended to be reproduced simultaneously with the CM service. In this case, the HI service shall be a single audio channel. This dialogue may have been processed for improved intelligibility by hearing impaired listeners. Simultaneous reproduction of both the CM and HI services allows the hearing impaired listener to hear a mix of the CM and HI services in order to emphasize the dialogue while still providing some music and effects.

Besides providing the HI service as a single dialogue channel, the HI service may be provided as a complete program mix containing music, effects, and dialogue with enhanced intelligibility. In this case, the service may be coded using any number of channels (up to 5.1). The fact that the service is a complete mix shall be indicated in the AC-3 descriptor (see Section 5.7.2.1 in Annex C of this Standard).

6.7 Dialogue (D)

The D associated service contains program dialogue intended for use with an ME main audio service. The language of the D service is indicated in the AC-3 bit stream, and in the audio descriptor. A complete audio program is formed by simultaneously decoding the D service and the ME service and mixing the D service into the center channel of the ME main service (with which it is associated).

If the ME main audio service contains more than two audio channels, the D service shall be monophonic (1/0 mode). If the main audio service contains two channels, the D service may also contain two channels (2/0 mode). In this case, a complete audio program is formed by simultaneously decoding the D service and the ME service, mixing the left channel of the ME service with the left channel of the D service, and mixing the right

channel of the ME service with the right channel of the D service. The result will be a two channel stereo signal containing music, effects, and dialogue.

Audio in multiple languages may be provided by supplying multiple D services (each in a different language) along with a single ME service. This is more efficient than providing multiple CM services, but, in the case of more than two audio channels in the ME service, requires that dialogue be restricted to the center channel.

6.8 Commentary (C)

The commentary associated service is similar to the D service, except that instead of conveying essential program dialogue, the C service conveys optional program commentary. The C service is always a single audio channel. When a C service(s) is provided, simultaneous decoding will give the listener the option to reproduce a C service along with a CM main service.

6.9 Emergency (E)

The E associated service is intended to allow the insertion of emergency or high priority announcements. The E service is always a single audio channel. An E service is given priority in transport and in audio decoding. Whenever the E service is present, it will be delivered to the audio decoder. Whenever the audio decoder receives an E type associated service, it will stop reproducing any main service being received and only reproduce the E service out of the center channel (or left and right channels if a center loudspeaker does not exits). The E service may also be used for non-emergency applications. It may be used whenever the broadcaster wishes to force all decoders to quit reproducing the main audio program and reproduce a higher priority single audio channel.

6.10 Voice-over (V0)

The VO associated service is a single channel service which is reproduced along with the main audio service in the receiver. It allows typical voice-overs to be added to an already encoded audio elementary stream without requiring the audio to be decoded back to baseband and then re-encoded. It is always a single audio channel. It has second priority (only the E service has higher priority). It is intended to be simultaneously decoded and mixed into the center channel of the main audio service. The dynamic range control signal in the VO service is intended to be used by the audio decoder to modify the level of the main audio program. Thus the level of the main audio service will be under the control of the broadcaster, and the broadcaster may signal the decoder (by altering the dynamic range control words embedded in the VO audio elementary stream) to reduce the level of the main audio service by up to 24 dB during the voice over.

7. AUDIO ENCODER INTERFACES

7.1 Audio encoder input characteristics

Audio signals which are input to the digital television system may be in analog or digital form. Audio signals should have any DC offset removed before being encoded. If the audio encoder does not include a DC blocking high pass filter, the audio signals should be high pass filtered before being applied to the encoder. In general, input signals should be quantized to at least 16-bit resolution. The audio compression system can convey audio signals with up to 24-bit resolution. Physical interfaces for the audio inputs to the encoder may be defined as voluntary industry standards by the AES, SMPTE, or other standards organizations.

7.2 Audio encoder output characteristics

Conceptually, the output of the audio encoder is an elementary stream which is formed into PES packets within the transport subsystem. It is possible that systems will be implemented wherein the formation of audio PES packets takes place within the audio encoder. In this case, the output(s) of the audio encoder(s) would be PES packets. Physical interfaces for these outputs (elementary streams and/or PES packets) may be defined as voluntary industry standards by SMPTE or other standards organizations.

Blank Page

ANNEX C

(Normative)

SERVICE MULTIPLEX AND TRANSPORT SYSTEMS CHARACTERISTICS

Table of Contents

List of Figures

List of Tables

ANNEX C

(Normative)

SERVICE MULTIPLEX AND TRANSPORT SYSTEMS CHARACTERISTICS

1. SCOPE

This Annex describes the transport layer characteristics and normative specifications of the Digital Television Standard.

2. NORMATIVE REFERENCES

The following documents contain provisions which in whole or in part, through reference in this text, constitute provisions of this Standard. At the time of publication, the editions indicated were valid. All standards are subject to revision and amendment, and parties to agreements based on this Standard are encouraged to investigate the possibility of applying the most recent editions of the documents listed below.

ATSC Standard A/52 (1994), *Digital Audio Compression (AC-3).*

ISO/IEC IS 13818-1, International Standard (1994), *MPEG-2 Systems.*

ISO/IEC IS 13818-2, International Standard (1994), *MPEG-2 Video.*

ISO/IEC CD 13818-4, MPEG Committee Draft (1994), *MPEG-2 Compliance.*

3. COMPLIANCE NOTATION

As used in this document, "*shall*" or "*will*" denotes a mandatory provision of the standard. "*Should*" denotes a provision that is recommended but not mandatory. "*May*" denotes a feature whose presence does not preclude compliance, that may or may not be present at the option of the implementor.

4. SYSTEM OVERVIEW

The transport format and protocol for the Digital Television Standard is a compatible subset of the MPEG-2 Systems specification defined in ISO/IEC 13818-1. It is based on a fixed-length packet transport stream approach which has been defined and optimized for digital television delivery applications.

As illustrated in Figure 1, the transport function resides between the application (e.g., audio or video) encoding and decoding functions and the transmission subsystem. The encoder's transport subsystem is responsible for formatting the coded elementary streams and multiplexing the different components of the program for transmission. At the receiver, it is responsible for recovering the elementary streams for the individual application decoders and for the corresponding error signaling. The transport subsystem

also incorporates other higher protocol layer functionality related to synchronization of the receiver.

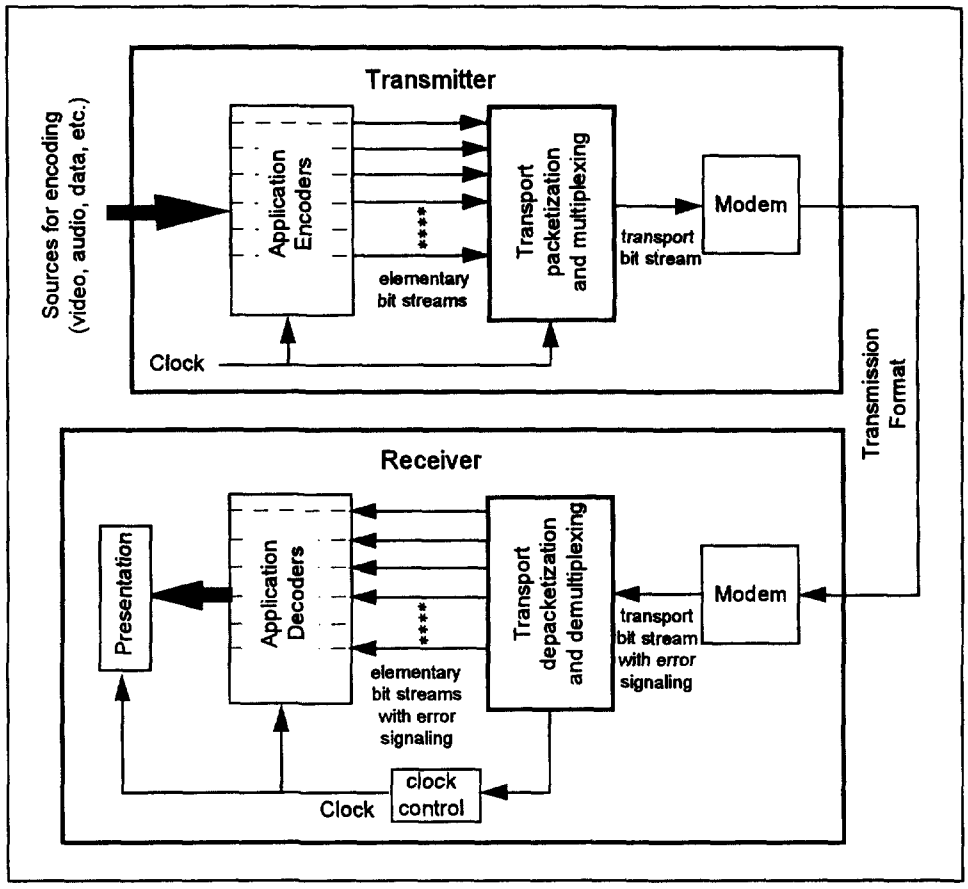

Figure 1. Sample organization of functionality in a transmitter-receiver pair for a single program.

The overall system multiplexing approach can be thought of as a combination of multiplexing at two different layers. In the first layer, single program transport bit streams are formed by multiplexing transport packets from one or more Packetized Elementary Stream (PES) sources. In the second layer, many single program transport bit streams are combined to form a system of programs. The Program Specific Information (PSI) streams contain the information relating to the identification of programs and the components of each program.

Not shown explicitly in Figure 1, but essential to the practical implementation of this Standard, is a control system that manages the transfer and processing of the elementary streams from the application encoders. The rules followed by this control system are not a part of this Standard but must be established as recommended practices by the users of the Standard. The control system implementation shall adhere to the requirements of the MPEG-2 transport system as specified in ISO/IEC 13818-1 with the additional constraints specified in this Standard. These constraints may go beyond the constraints imposed by the application encoders.

5. SPECIFICATION

This Section constitutes the normative specification for the transport system of the Digital Television Standard. The syntax and semantics of the specification conform to ISO/IEC 13818-1 subject to the constraints and conditions specified in this Standard. This Section of the Standard describes the coding constraints that apply to the use of the MPEG-2 systems specification in the digital television system.

5.1 MPEG-2 Systems standard

The transport system is based on the transport stream definition of the MPEG-2 Systems standard as specified in ISO/IEC 13818-1.

5.1.1 Video T-STD

The video T-STD follows the constraints for the "High Level" in Section 2.4.2.3 of ISO/IEC 13818-1.

5.1.2 Audio T-STD

The audio T-STD is specified in Section 3.6 of Annex A of ATSC Standard A/52.

5.2 Registration descriptor

This Standard uses the registration descriptor described in Section 2.6.8 of ISO/IEC 13818-1 to identify the contents of programs and elementary streams to decoding equipment.

5.2.1 Program identifier

Programs which conform to this specification will be identified by the 32-bit identifier in the section of the Program Map Table (PMT) detailed in Section 2.4.4.8 of ISO/IEC 13818-1. The identifier will be coded according to Section 2.6.8, and shall have a value of 0x4741 3934.

5.2.2 Audio elementary stream identifier

Audio elementary streams which conform to this specification will be identified by the 32-bit identifier in the section of the Program Map Table (PMT) detailed in Section 2.4.4.8 of ISO/IEC 13818-1. The identifier will be coded according to Section 2.6.8, and shall have a value of 0x4143 2D33.

5.3 The program paradigm

The program paradigm specifies the method that shall be used for allocating the values of the Packet Identifier (PID) field of the transport packet header in a systematic manner. Within one transport multiplex, television programs that follow the program paradigm are assigned a program number ranging from 1 to 255. The binary value of the program number is used to form b_{11} through b_4 of the PID. Programs adhering to the

paradigm shall have b_{12} equal to '0'. Programs not adhering to the paradigm shall have b_{12} equal to '1'.

We further define:

- base_PID = program number << 4

The b_0 through b_3 of the PID are assigned according to Table 1.

The paradigm to identify the transport bit streams containing certain elements of the program is defined in Table 1.

Table 1 PID Assignment for the Constituent Elementary Streams of a Program

Name	PID Definition	Description
PMT_PID	base_PID+0x0000	PID for the bit stream containing the program_map_table for the program.
Video_PID	base_PID+0x0001	PID for the bit stream containing the video for the program.
PCR_PID	base_PID+0x0001	Implies the video bit stream also carries the PCR values for the program
Audio(A)_PID	base_PID+0x0004	PID for the bit stream containing the primary audio for the program. This may be either a complete main audio service (CM) as defined in Section 6.3 of Annex B or a main audio service, music and effects (ME) as defined in Section 6.4 of Annex B.
Audio(B)_PID	base_PID+0x0005	When the Audio (A) bit stream contains a complete main audio service (CM), the Audio(B)_PID shall not appear in the transport stream. When the Audio (A) bit stream contains a main audio service, music and effects (ME), the Audio (B) bit stream contains the principle program dialogue (D) associated with the ME service carried by Audio (A).
Data_PID	base_PID+0x000A	PID for the bit stream containing the data for the program.

The program_map_table must be decoded to obtain the PIDs for services not defined by the paradigm but included within the program (such as a second data channel). According to the program paradigm, every 16th PID is a PMT_PID and may be assigned to a program. If a PMT_PID is assigned to a program by the program paradigm, the next 15 PIDs after that PMT_PID are reserved for elements of that program and shall not be otherwise assigned.

5.4 Constraints on PSI

The program constituents for all programs, including television programs that follow the program paradigm and other programs or services that do not follow the program paradigm, are described in the PSI. There are the following constraints on the PSI information:

- Only one program is described in a PSI transport bit stream corresponding to a particular PMT_PID value. A transport bit stream containing a program_map_table shall not be used to transmit any other kind of PSI table (identified by a different table_id).